Published in final edited form as:

Environmetrics. 2018 February ; 29(1): . doi:10.1002/env.2477.

Zero-inflated multiscale models for aggregated small area health data

Mehreteab Aregay^{1,*}, Andrew B Lawson¹, Christel Faes², Russell S Kirby³, Rachel Carroll⁴, and Kevin Watjou²

¹Department of Public Health, Medical University of South Carolina, Charleston SC USA

²Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium

³Department of Community and Family Health, University of South Florida, Tampa, FL USA

⁴Biostatistics & Computational Biology Branch National Institute of Environmental Health Sciences, Durham NC USA

Abstract

It is our primary focus to study the spatial distribution of disease incidence at different geographical levels. Often, spatial data are available in the form of aggregation at multiple scale levels such as census tract, county, state, and so on. When data are aggregated from a fine (e.g. county) to a coarse (e.g. state) geographical level, there will be loss of information. The problem is more challenging when excessive zeros are available at the fine level. After data aggregation, the excessive zeros at the fine level will be reduced at the coarse level. If we ignore the zero inflation and the aggregation effect, we could get inconsistent risk estimates at the fine and coarse levels. Hence, in this paper, we address those problems using zero inflated multiscale models that jointly describe the risk variations at different geographical levels. For the excessive zeros at the fine level, we use a zero inflated convolution model, whereas we consider a regular convolution model for the smoothed data at the coarse level. These methods provide a consistent risk estimate at the fine and coarse levels when high percentages of structural zeros are present in the data.

Keywords

Structural zeros; sampling zeros; zero inflated models; multiscale models; scaling effects

1 Introduction

It is natural to collect data at different scale levels such as census tract, county, and state. Research shows that aggregating data from a fine (e.g. county) to a coarse (e.g. state) geographical level results in a loss of information at the coarse level [1–6]. For example, there is a loss of information when we average county-level data to obtain the data at the state level. This process results in smoothing out the data at the state level is known as

*Corresponding author, mhrete_ab@yahoo.com.

scaling effects. The problem is more challenging when there are excessive zeros at the fine level because the excessive zeros at this level will be reduced at the coarse level during data aggregation.

A prime example of excessive zeros in spatial data is often observed for rare cancers or low incidence cancers in a particular area or a particular period. When count data consist of a large proportion of zero values, the standard Poisson model could fit the data poorly because it assumes the conditional mean is equal to the conditional variance. For the zero inflated count data, the conditional variance is often much larger than the conditional mean, a feature known as overdispersion. To solve such problem in spatial data, researchers have implemented zero inflated Poisson (ZIP) models and also hurdle models [7–10].

The difference between these two approaches is that the ZIP model assumes a mixture of Bernoulli and Poisson distributions for the response variable, whereas the hurdle model assumes a Bernoulli distribution and a truncated-at-zero Poisson distribution for the zero and positive counts, respectively [11–12]. Moreover, the ZIP model considers two kinds of zeros, i.e., structural and sampling zeros, while the hurdle model employs only structural zeros. Structural zeros represent a subgroup of individuals who are not at risk for certain diseases, while sampling zeros are present for at-risk groups [13]. In particular, structural zeros occur because of a separate Bernoulli process, whereas sampling zeros arise from a Poisson distribution. Hurdle models are often used when the structural zeros are known. However, when the distinction between structural and random zeros is unknown, ZIP is the natural choice. Since we do not have information whether the structural zeros are known in our example in Section 2, in this paper, we consider ZIP models.

Although there is much research on spatial zero inflated models, limited examples are available on extending ZIP models to multiscale data [14]. On the other hand, many studies have been done on multiscale models to address scaling effects [15–19] but these models have not been applied to zero inflated data. Hence, our aim is to accommodate both scaling effect as well as zero inflated problems using a general, simple, and practical, yet flexible enough spatial multiscale model for zero inflated multiscale data. If we ignore the excessive zeros and the scaling effects, the consequence could be significant. For instance, it can result in an inconsistent disease incidence estimate at the fine and coarse levels, i.e., we may obtain an elevated risk at the county level, whereas we may find a diminished risk at the coarse level. Jointly modeling the risk variation at the fine and coarse levels is very useful and it has the following attractive features. 1) It allows one to obtain a more accurate risk estimate at both the fine and coarse levels in a single analysis; 2) it permits the incorporation of the effect of a parent (e.g. state) on the children (e.g. counties); and 3) it provides a consistent result at both the fine and coarse levels. Our method is unique in the sense that we only use a zero inflated model for the data at the fine level. We use a standard method at the coarse level since the excess of zeros will be reduced during data aggregation. Note, however, that if many zeros exist at the coarse level, these models can be easily extended to accommodate the zero inflation at the coarse level as well.

In particular, we consider four different models. The first two models (M1 and M2) are zero inflated models at the fine level with and without shared components that jointly link the

data across scales. Our focus is to compare M1 versus M2 via an extensive simulation study as well as in a real life application. As a secondary aim, we also consider two other standard multiscale models (M3 and M4) with and without the shared components. These models do not contain extra-parameters that handle the excessive zeros in the multiscale data. However, M3 and M4 consist of correlated and uncorrelated random effects. Besides assessing the effect of zero inflation with aggregation on estimation accuracy, we are interested in assessing whether these random effects in M3 and M4 can address overdispersion due to excessive zeros.

The paper is organized as follows. Section 2 describes the skin cancer example from the state of Georgia. Section 3 details the statistical methods and design of the simulation study, while Section 4 presents the results from the models applied to the real and simulated data sets. Finally, Section 5 provides a discussion and concluding remarks.

2 Georgia Skin Cancer Study

We examine skin cancer data available from the state of Georgia via the Georgia Division of Public Health OASIS system (<http://oasis.state.ga.us>) at both the county and public health (PH) district levels in 2008. In Georgia, 159 counties are aggregated into 18 public health districts. The aim of organizing the counties to PH districts is to improve the health of the people who live in the counties within the PH districts. This is an aligned aggregation (see Figure 1) meaning that the counties belong within the given PH district. It is an example of a modifiable areal unit problem (MAUP; [20–21]) where the scaling problem is present due to aligned data aggregation. Here, we do not address a spatial misalignment problem [22–23]. Notice that ecological bias is not a concern here because we make inference at the higher aggregation level. Ecological bias should be considered when inference is to be made from an aggregated study to an individual level ([2, ch 9]). However, the atomistic fallacy should be a concern because we are inferring from a lower into a higher geographical level. We attempt to address the atomistic fallacy using the shared components that jointly link the data across scales.

Our interest lies in investigating whether the incidence of skin cancer is spatially related to daily sunlight (KJ/m^2) attained from the North America Land Data Assimilation System [24]. Figure 1 displays the average daily sunlight (ADS) in thousands of KJ/m^2 as well as the standardized incidence risk (SIR) computed as a ratio of the observed response to the expected number of cases. At both county and PH levels, the incidence of skin cancer is positively related with daily sunlight. For example, higher values of SIR and ADS are present in the coastal health district of Georgia. In some of the counties within the coastal health district such as McIntosh, the incidence of skin cancer and ADS are high. The ADS increases from North to South. Hence, it has a spatial pattern. Note that in the Northwest and southwest of Georgia, the SIR is zero for most of neighbored counties. This shows that there are clusters of zero risks meaning that neighbored counties in those regions tend to have zero risks. In this example, the maximum number of skin cancer cases at the county level in 2008 is 6, while 122 counties (76.7%) have zero cases (Figure 2). After aggregation, however, the percentage of zeros decreases to 16% at the PH district level (higher-level health administrative units). This shows that zero inflation could be present at the county-level,

whereas scaling effects exist at the PH level. Hence, we aim to address these problems using models that are in line with the nature of the data. Note that the 76.7% zeros at the county level could come from the sampling and structural zeros.

3 Methodology

The proposed zero inflated multiscale models use joint convolution models that describe the risk variation at multiple scale levels via a shared random effect component. Our methods have the following advantages. First, they are based on flexible joint models, which accommodate the hierarchical structure of multiscale data. Second, they allow a more complicated model to be considered for the zero skewed data at the fine level while it continues to benefit from using a relatively parsimonious model for the spatial ordinary count data at the coarse level. In the next section, we describe our model formulation for heavily zero inflated data at two scale levels. The models can be easily extended beyond two scale levels. For simplicity, we called the Models 1-4 below as M1, M2, M3, and M4.

3.1 Zero inflated shared multiscale model (M1)

Our zero inflated multiscale models extend the convolution model proposed by Besag et al [25]. It is notable that the convolution model consists of correlated and uncorrelated heterogeneity. In this paper, the fine level zero inflated convolution model is connected with the standard coarse-level convolution model via a shared random effect. The linkage component could be one of the following: (1) a correlated heterogeneity (CH), (2) an uncorrelated heterogeneity (UH), or (3) both the CH and the UH. If the correlated heterogeneity predominates the uncorrelated heterogeneity at the coarse level, it is reasonable to use the correlated heterogeneity as a linkage component because the CH will explain the aggregation effect better than the UH. Otherwise, the UH component could be used as a linkage component. On the other hand, if both the CH and UH equally explain the aggregation effect, we can use both CH and UH to jointly link the data at each scale. As an example, we demonstrate the multiscale convolution models by sharing both the CH and the UH for skin cancer in the state of Georgia. Assume y_{ij} is the number of persons discharged from non-federal acute-care inpatient facilities for skin cancer in 2007 for the j^{th} county within the i^{th} PH district, where $j=1, 2, \dots, M_i$ and $i=1, 2, \dots, 18$. Here, M_i represents the number of counties within the i^{th} PH district. The zero inflated Poisson model at the county level has the following form:

$$P(Y_{ij}=y_{ij}|z_{ij}, x_{ij}) = \begin{cases} p_{ij} + (1 - p_{ij})Pois(\mu_{ij}; 0|x_{ij}) & \text{if } y_{ij} = 0 \\ (1 - p_{ij})Pois(\mu_{ij}; y_{ij}|x_{ij}) & \text{if } y_{ij} > 0, \end{cases} \quad (1)$$

To model the probability p_{ij} , a logit link, $logit(p_{ij}) = z'_{ij}\alpha$, is often used whereas a log link function, $\log(\mu_{ij}) = x'_{ij}\beta$, is applied to link μ_{ij} with a linear predictor; x'_{ij} and z'_{ij} are the vector of covariates defining the mean μ_{ij} and the probability p_{ij} , respectively [26]. Further, $\beta' = (\beta_1, \beta_2, \dots, \beta_p)$ and $\alpha' = (\alpha_1, \alpha_2, \dots, \alpha_p)$ are the vector of unknown parameters. We assume a non-informative beta distribution, $beta(1, 1)$ for p_{ij} . We use a linear predictor only

for the Poisson outcome. Hence, we focus on whether the risk of skin cancer is spatially related to daily sunlight (KJ/m^2). Moreover, we assume $\mu_{ij} = e_{ij}\theta_{ij}$, where e_{ij} is the expected number of cases and θ_{ij} is the relative risk for the j^{th} county within the i^{th} PH district. In

addition, e_{ij} is given by:
$$e_{ij} = \frac{\sum_{i=1}^{18} \sum_{j=1}^{M_i} y_{ij}}{\sum_{i=1}^{18} \sum_{j=1}^{M_i} \pi_{ij}} \pi_{ij}$$
 where π_{ij} is the county level population size in county j within the PH district i . For the example in Section 2, θ_{ij} is given by:

$$\log(\theta_{ij}) = a_{01} + \beta_{11} x_{ij} + v_{ij} + u_{ij} + v_i + u_i, \quad (2)$$

where a_{01} , β_{11} , v_{ij} , and u_{ij} are the intercept, slope, UH, and CH at the county level, while v_i and u_i are the shared UH and CH at the PH district level, respectively. Hence, v_i and u_i are the same for all the counties within the PH district meaning that the children (counties) will share the common characteristics, which they inherit from their parent (PH district). Here, x_{ij} denotes the average daily sunlight (ADS) in the j^{th} county within the i^{th} PH district. To improve convergence, we standardize the covariate ADS. We consider Gaussian distributions for the uncorrelated components, whereas we use an intrinsic conditional autoregressive distributions (ICAR) for the correlated components as follows: $v_{ij} \sim N(0, \sigma_{v1}^2)$, $v_i \sim N(0, \sigma_{v2}^2)$, $u_{ij} \sim \text{ICAR}(\sigma_{u1}^2)$, and $u_i \sim \text{ICAR}(\sigma_{u2}^2)$. We define correlation based on adjacent areas. For

example, the PH-district ICAR can be expressed as: $u_i \sim N\left(\frac{1}{n_i} \sum_{i \sim l} u_l, \frac{\sigma_{u2}^2}{n_i}\right)$, where $i \sim l$ denotes PH district i is adjacent to PH district l and n_i is the number of PH districts that are adjacent to the i^{th} PH district.

So far, we have seen how to model the excessive zeros at the county level. The data at the PH levels are obtained by aggregating the data at the county level, i.e., $y_i = \sum_{j=1}^{M_i} y_{ij}$ and $x_i = \sum_{j=1}^{M_i} x_{ij} / M_i$ with y_i and x_i are the number of skin cancer cases and the ADS at PH district i , respectively. Because of scaling effects, the excessive zeros at the county level are smoothed out at the PH level and we have only 16% of excessive zeros at the PH district. Hence, we assume that y_i follows a conditionally independent Poisson distribution, i.e., $y_i | \theta_i \sim \text{Pois}(\mu_i = e_i \theta_i)$ where e_i is the expected number of cases in PH district i and θ_i is the PH-level relative risk in the form

$$\log(\theta_i) = a_{02} + \beta_{12} x_i + v_i + u_i \quad (3)$$

Here a_{02} and β_{12} are the intercept and slope parameters. Note that the shared components v_i and u_i link (2) and (3). For this and the subsequent models, we use non-informative normal prior distributions for the “fixed” effect parameters, whereas we employ uniform prior distributions for the standard deviations of the random effects [27]. In particular, we considered non-informative normal priors for $\beta_* = (a_{01}, a_{02}, \beta_{11}, \beta_{12})$, i.e.,

$\beta_* \sim N(0, \sigma_\beta = 100)$ and uniform priors for the standard deviation of the random effects, $\sigma_* = (\sigma_{u1}, \sigma_{v1}, \sigma_{u2}, \sigma_{v2})$, i.e., $\sigma_* \sim Unif(0, 100)$.

We fitted all models jointly using Markov chain Monte Carlo (MCMC) posterior sampling in a single analysis. We considered 15,000 iterations after we discarded the first 15,000 burn-in samples. We ran three separate chains starting from different initial values for each model. Hence, the posterior means were calculated based on 45,000 iterations, which were sufficient for convergence. Further, we have run the MCMC algorithm for 100,000 burn-in iterations followed by 15,000 iterations and we found the same estimates for the parameters as when we ran for 15,000 and then sampled 15,000. We assessed convergence using an estimated potential scale reduction factor ($\hat{R}$) and trace plots.

3.2 Zero inflated independent multiscale model (M2)

Here, we are interested in assessing the impact of ignoring the scale effects for aggregated small area health data with heavily skewed zeros. This model ignores the scaling effects but it still accommodates for the excessive zeros at the county level. Hence, M2 can be expressed as in (1–3) without the shared components in (2), which is replaced by

$$\log(\theta_{ij}) = a_{01} + \beta_{11}x_{ij} + v_{ij} + u_{ij} \quad (4)$$

The convolution model at the PH level is the same as in (3). Note that there is no linkage component that jointly connects the relative risk (RR) of the zero inflated model in (4) to the RR of the convolution model in (3).

3.3 Shared multiscale model (M3)

This model ignores the excessive zeros at the county level and uses convolution models at both the county and PH levels. Hence, we assume a conditionally Poisson distribution at both levels, i.e., $y_{ij}|\theta_{ij} \sim Pois(\mu_{ij} = e_{ij}\theta_{ij})$ and $y_i|\theta_i \sim Pois(\mu_i = e_i\theta_i)$. The county-level θ_{ij} is the same as in (2), whereas the PH-level θ_i is the same as in (3). Hence, the relative risks in this model can be expressed as:

$$\log(\theta_{ij}) = a_{01} + \beta_{11}x_{ij} + v_{ij} + u_{ij} + v_i + u_i, \quad \log(\theta_i) = a_{02} + \beta_{12}x_i + v_i + u_i \quad (5)$$

Here, we are interested in investigating whether the CH and the UH terms, i.e., v_{ij} , u_{ij} , v_i and u_i in (5) could address the overdispersion due to excessive zeros.

3.4 Independent multiscale model (M4)

M4 is the most parsimonious model and it is the simplified version of M3. It neither includes an additional mixing probability to address the excessive zeros nor shared components to account for the scaling effects. However, the CH and the UH in this model as well as in M3 may handle some amount of the overdispersion caused by the excessive zeros. Similar to M3, here also, we assume a conditional Poisson distribution at both levels and the relative risks are of the form:

$$\log(\theta_{ij}) = a_{01} + \beta_{11}x_{ij} + v_{ij} + u_{ij}$$

$$\log(\theta_i) = a_{02} + \beta_{12}x_i + v_i + u_i$$

3.5 Model Assessment and Goodness of Fit

To select the best model, we consider a range of information criteria. First, we employ the deviance information criterion (DIC₃) developed by Celeux et al. [28] for mixtures of distributions. Celeux et al study different possible variations of the DIC proposed by Spiegelhalter et al. [29]. They find that DIC₃ and DIC₄ were the most reliable tools as compared to the other eight criteria, which they considered. Here, we use DIC₃ for ease of implementation in widely available software such as WinBUGS. In addition, DIC₃ is a sensible alternative because it uses predictive distribution, which is central to the Bayesian inference [28]. It can be expressed as:

$$\text{DIC}_3 = -4E_\theta[\log f(\mathbf{y}|\boldsymbol{\theta})|\mathbf{y}] + 2\log \hat{f}(\mathbf{y}),$$

where $f(\mathbf{y}|\boldsymbol{\theta})$ is the observed likelihood function and $\hat{f}(\mathbf{y}) = \prod_{i=1}^n \hat{f}(y_i)$. Second, we consider DIC, which is composed of deviance and the effective number of parameters (p_D). Since the DIC proposed by Spiegelhalter et al. [29] can lead to a negative result for p_D in the mixture case, we employ the p_D proposed by Gelman et al. [30]. Hence, this DIC is provided as:

$$\text{DIC}_G = \overline{D(\boldsymbol{\theta})} + p_D,$$

where $p_D = \text{var}(D(\boldsymbol{\theta}))/2$, with $\text{var}(D(\boldsymbol{\theta}))$ is the estimated posterior variance of the deviance and $\overline{D(\boldsymbol{\theta})} = E_\theta[-2 \log f(\mathbf{y}|\boldsymbol{\theta})|\mathbf{y}]$.

Furthermore, we consider a marginal predictive likelihood (MPL) to select the best model. For model M , MPL can be expressed as follows:

$$f(\mathbf{y}|M) = \int f(\mathbf{y}|\boldsymbol{\theta}, M) f(\boldsymbol{\theta}|M) d\boldsymbol{\theta} \quad (6)$$

As we can see in (6), the MPL is an expectation over the likelihood $f(\mathbf{y}|\boldsymbol{\theta}, M)$ with respect to the prior distribution of the parameters $f(\boldsymbol{\theta}|M)$ [31].

In this paper, we aim to compare M1 with M2 to assess the benefit of including a shared component in the zero inflated models. We are also interested in comparing the zero inflated multiscale models (M1 and M2) with multiscale models (M3 and M4). In particular, we compare M1 versus M3 and M2 versus M4 to investigate the important of addressing the county-level overdispersion due to excessive zeros. Since we study M3 with only CH shared

component and M4 in our previous paper [3], we do not focus on the comparison between M3 and M4 here.

3.6 Simulation Study

The simulation study is designed to evaluate Models 1-4 for their ability to address the following issues that could be encountered in a real life application. 1) Clusters of zero risks: in a rare disease, the RR in neighboring areas may be zero. Hence, in our simulation, we fix the RR as zero for certain neighboring areas and evaluated whether the models could recover the simulated risks. 2) Structural zeros: here, we are interested in assessing the impact of ignoring high percentage of structural zeros such as 25% and 50%. 3) Scaling effect: in this scenario, we simulate first PH-level data from a Poisson distribution. Thereafter, we simulate county-level data from a multinomial distribution by conditioning on the PH-level data. We explain each of these scenarios in the next section.

3.6.1 Scenario 1: Clusters of zero risks

3.6.1.1 Scenario 1.1: Simulation from M3: First, we simulate the county-level data y_{ij} , similar to M3, from a conditional independent Poisson distribution, $y_{ij}|\theta_{ij} \sim \text{Pois}(\mu_{ij} = e_{ij}\theta_{ij})$, where θ_{ij} as in (2) and e_{ij} is simulated from a gamma distribution with hyper parameters equal to 1, $e_{ij} \sim \text{Gamma}(1, 1)$. To sample θ_{ij} , the following values are assumed for the parameters in (2): $a_{01} = \beta_{11} = 0.1$, and $\sigma_{v1} = \sigma_{u1} = \sigma_{v2} = \sigma_{u2} = 0.3$. Thereafter, the uncorrelated random effects are simulated from a normal distribution, whereas the correlated random effects are generated from the ICAR distribution as follows:

$v_{ij} \sim N(0, \sigma_{v1} = 0.3)$, $v_i \sim N(0, \sigma_{v2} = 0.3)$, $u_{ij} \sim \text{ICAR}(\sigma_{u1} = 0.3)$, and $u_i \sim \text{ICAR}(\sigma_{u2} = 0.3)$. We sample x_{ij} from a standard normal distribution, $x_{ij} \sim N(0, 1)$. In order to create clusters of zero risks, we fix θ_{ij} to be equal to zero for some neighbored areas. The simulated county-level θ_{ij} is shown in the left panel of Figure 3, indicating that the RRs in some neighboring counties in south, central, and northern Georgia are zeros. The PH-level relative risks, θ_i , in the right panel are computed by aggregating the county-level RRs, $\theta_i = \mu_i / e_i$,

$\mu_i = \sum_{j=1}^{M_i} \theta_{ij} e_{ij}$, with M_i and e_i as described in Section 3.1. Similarly, the PH-level outcome

is obtained by summing up the simulated county-level outcomes, $y_i = \sum_{j=1}^{M_i} y_{ij}$ and the PH-level covariate is computed by averaging the county-level predictor values,

$x_i = \sum_{j=1}^{M_i} x_{ij} / M_i$. In this and subsequent scenarios, we simulate 200 data sets and we fit and compare Models 1-4 by averaging the DICs values, bias, and MSE estimates of the parameters over the 200 simulated data sets. As an example, we present in Figure 1a (see in the Supplementary appendix) one realization of the simulated outcome at the county and PH levels. For the county level, the percentage of zeros are 52%, while at the PH level, they are 17%. This indicates that the excessive zeros at the county level are smoothed out when they are aggregated to the PH-level due to the scaling effect. Note that, although we simulate the county-level data from M3, which ignores the excessive zeros, we introduce excess zeros to the data by fixing some counties to have zero risks.

3.6.1.2 Scenario 1.2: Simulation from Equation 4: This simulation design is similar to scenario 1.1, except, we simulate the county-level RRs, θ_{ij} , as in (4). This design can help us to assess whether the complex models, M1-M3, simplify to M4. We make the following assumptions about the parameters in (4): $a_{01}=\beta_{11}=0.1$, and $\sigma_{v1}=\sigma_{u1}=0.3$. The random effects, v_{ij} and u_{ij} , and the predictor x_{ij} are generated in a similar fashion as in scenario 1.1. Here also, we fix the RRs equal to zero for some neighboring areas as shown in Figure 4. Thereafter, we simulate y_{ij} from a Poisson distribution, $y_{ij}|\theta_{ij}\sim\text{Pois}(\mu_{ij}=e_{ij}\theta_{ij})$, with $e_{ij}\sim\text{Gamma}(1, 1)$. y_i , x_i , and θ_i are computed as in scenario 1.1.

3.6.1.3 Scenario 1.3: Simulation from a Poisson Model: One can argue that the CH and the UH in M3 and M4 can address to a degree overdispersion due to excessive zeros. Hence, in this scenario, we simulate y_{ij} with excessive zeros from the following simple Poisson model, i.e., $y_{ij}|\theta_{ij}\sim\text{Pois}(\mu_{ij}=e_{ij}\theta_{ij})$, where $\log(\theta_{ij})=a_{01}+\beta_{11}x_{ij}$ with $e_{ij}\sim\text{Gamma}(1, 1)$. As in scenario 1.1., we assume $a_{01}=\beta_{11}=0.1$, $x_{ij}\sim N(0, 1)$ as well as we fix θ_{ij} equals to 0 for some neighboring counties. The simulated RRs are shown in Supplementary Figure 2A. We simulate 200 data sets and we fit Model 1A (M1A) and Model 2A (M2A) without the random effects as shown in Supplementary Table 1A. Note that M1A is a simplified version of M1 and M2, whereas M2A is the parsimonious model for M3 and M4. We also apply M1-M4 to the simulated data to investigate how these complex models behave when the data are generated from a simple Poisson distribution with excessive zeros.

3.6.2 Scenario 2: Simulation from a ZIP Model—In the previous scenarios 1.1-1.3, we fix the counties with zero risks, and then, we generate data from a Poisson distribution at the county level. In this scenario, first, we simulate the relative risks θ_{ij} from (2) as shown in the Supplementary Figure 3A. Note that the assumed values of the model parameters and e_{ij} are the same as in scenario 1.1. Thereafter, we simulate 200 data sets from the ZIP model in (1) at the county level using the *rzipois* function available in **VGAM** package in R [32]. We consider 0%, 10%, 25% and 50% structural zeros (i.e. ignoring the Poisson distribution). Supplementary Figure 4A displays an example of the simulated data set with 25% structural zeros at the county level. In this example, we have a total of 57% zeros in which the 25% represent the structural zeros (non-risk groups), whereas 32% represent random (sampling) zeros, i.e., they are part of the Poisson data. Poisson models are not flexible enough to accommodate the difference between structural and random zeros. For a detailed description about the difference between structural and sampling zeros, we refer to He et al [13]. As in scenarios 1.1-1.3, the data at the PH level are an aggregation of the county-level data.

3.6.3 Scenario 3: Scaling effect—In scenario 1, we only simulate the county-level outcome and then we obtain the PH-level data by summing up the county-level data within the PH district. Although this approach introduces a scaling effect due to data aggregation, it does not provide a tool to investigate whether the parameters at the PH-levels in M1-M4 are unbiased and precise estimates. Hence, in scenario 3, we simulate data at both the county and PH levels. First, we simulate PH-level data from a Poisson distribution as follows: $y_i|\theta_i\sim\text{Pois}(\mu_i=e_i\theta_i)$, where $\log(\theta_i)=a_{02}+\beta_{12}x_i+v_i+u_i$. To focus on the parameter of interest θ_i , we assume the expected rate equals 1, $e_i=1$ for all i . We make the following

assumptions for the model parameters: $a_{02}=\beta_{12}=0.5$, $v_i \sim N(0, \sigma_{v2}=0.3)$, $u_i \sim ICAR(\sigma_{u2}=0.3)$. Conditioning on the simulated PH-level data, we sample the county-level data from a multinomial distribution as following: $y_{ij}|(p_{ij}, y_i) \sim Multinomial(y_i, p_{ij})$,

where $p_{ij} = \frac{e_{ij}\theta_{ij}}{\sum_{j=1}^{M_i} e_{ij}\theta_{ij}}$ with $\log(\theta_{ij}) = a_{01} + \beta_{11}x_{ij} + v_{ij} + u_{ij} + v_i + u_i$, and $e_{ij}=1$ for all i and j .

Here also, we assume the following values: $a_{02}=\beta_{12}=0.5$, $v_i \sim N(0, \sigma_{v2}=0.3)$, $u_i \sim ICAR(\sigma_{u2}=0.3)$, $v_{ij} \sim N(0, \sigma_{v1}=0.3)$, $u_{ij} \sim ICAR(\sigma_{u1}=0.3)$, $x_{ij} \sim N(0, 1)$. Note that we

average x_{ij} , to obtain x_i as follows: $x_i = \sum_{j=1}^{M_i} x_{ij} / M_i$. For this scenario, the simulated risks are shown in Figure 5. We also present one realization of the 200 simulated data sets in Supplementary Figure 5A indicating that there are nearly 72% of zeros at the county-level data, whereas there are 0% of zeros at the PH-level data.

4. Results

4.1 Simulation Results

4.1.1 Scenario 1.1: Simulation from M3—Table 1 displays the goodness of fit (GoF) measures. The DIC_3 and MPL values suggest that M4 is the best model at the county level, whereas M3 is the best model at the PH-level. However, DIC_G indicates that M1 is the best model at the county level, while M3 is still the best model at the PH level. For this scenario, the expected best fit is M1 at the county-level because the data are simulated with excessive zeros from a Poisson distribution with mean structure of θ_{ij} as in (2). Hence, DIC_G provides the expected result, while DIC_3 and MPL produce unexpected results. At the PH level, we anticipate M1 and M3 to provide a better fit because the shared components could address the scaling effect due to data aggregation; all the criteria confirm this expectation. However, all the criteria prefer M3 as compared to M1 at the PH level. We find similar results in scenarios 1.2, 1.3, and 2 below. Note that M1 and M3 have the same model formulation at the PH level but they are different at the county level. The same is true for M2 and M4. When we compare the zero inflated models (M1 versus M2), M1 outperforms M2 across the levels. This shows that the shared components in M1 are useful to link the risks across the levels.

With respect to the parameter estimates, as expected, M1 followed by M2 offers the most unbiased and precise estimates. However, M3 and M4 results in more biased and imprecise estimates suggesting that ignoring the overdispersion due to excessive zeros may yield biased and imprecise results (Table 2). Note that M3 and M4 underestimate the overall relative risk, ($e^{a_{01}}$), while they tend to overestimate the covariate effect (β_{11}), contextual effect (σ_{u2} and σ_{v2}), and county-level variance components (σ_{u1} and σ_{v1}).

In terms of recovering the simulated risks, M3 and M4 recover the clusters of zero risks in Figure 3 better than M1 and M2 (Figure 6), especially at the county level. However, M1 detects the elevated risks to northeastern Georgia better than the other models. This is an important result because in public health, we are often interested in identifying high-risk areas. Moreover, at the PH level, M1 recovers the simulated risks well when compared to the

other models. Note that the range of the county-level simulated risks in Figure 3 is 0 to 3.09, while the range of the county-level risks obtained from M1, M2, M3, and M4 in Figure 6 are 0.34 to 3.11, 0.97 to 2.34, 0.13 to 2.64, and 0.17 to 2.54, respectively.

4.1.2 Scenario 1.2: Simulation from M4—Similar to scenario 1.1, M1 is the best model at the county level using DIC_G measures followed by M2. Yet, M3 is the best model at the county level using DIC_3 . MPL slightly prefers M4 as compared to M3 at the county level. However, at the PH level, all the three criteria opt for M3 followed by M1, indicating that the shared components in M1 and M3 improve the GoF at the PH level. We expect M1 and M2 to be the best models as the data sets are generated with excessive zeros at the county level and DIC_G offers the expected results. At the PH level, we anticipate M1 and M3 to outperform M2 and M4 and all the criteria confirm our expectation (Table 3). As in scenario 1.1, M1 is better than M2.

As for the estimation of the parameters, M1 yields the most unbiased and precise estimates for most of the parameters. Similar to scenario 1.1, M2 provides the most unbiased and precise estimates for the intercept term. However, M3 and M4 result in larger bias and MSE estimates for most of the parameters (Table 4).

With respect to recovering the simulated risks, M3 recovers the county-level risk pattern, especially the clusters of zero risks, better than the other models. Although M1 and M2 slightly overestimate the cluster of zero risks, they correctly identify the elevated risks for some counties better than M3 and M4. For example, in southwest Georgia, M1 and M2 offers higher county-level risks than M3 and M4. At the PH level, all the models provide similar risk estimates. However, M1 slightly underestimates the PH-level risks in the southwest and southeast of Georgia when compared to the other models (Figure 7).

4.1.3 Scenario 1.3: Simulation from a Poisson Model—The results obtained from M1-M4 are shown in Supplementary Tables 2A and 3A. The findings from this scenario are similar with the findings in scenarios 1.1 and 1.2. Here also, DIC_3 picks M3, whereas DIC_G selects M1 as the best model at the county level. However, when we compare M1A and M2A (see Supplementary Table 4A), both DIC_3 and DIC_G agree and pick M1A as the best county-level model. This is interesting because both criteria, unlike in Models 1-4, agree and provide an expected result. Note that, unlike M3 and M4, M2A does not contain random effects that could adjust for overdispersion effects. Hence, it lacks enough parameters to fit the simulated data adequately. On the other hand, M3 and M4 consist of correlated and uncorrelated random effects, which could handle to some extent an overdispersion due to excessive zeros. This could be the reason why DIC_3 selects M3 and M4 over M1 and M2 although we simulate the data with excessive zeros. We also notice that MPL opts for M1A as the county-level best model.

When we compare M1 versus M2, as in scenarios 1.1 and 1.2, M1 outperforms M2 across the range of criteria. This shows that even when the data are simulated from a simple ZIP model without random effects, the shared components in M1 are useful to describe jointly the risk variations across the scales.

Another interesting result is that M2 and M1A perform similarly at the county level. This is an expected result because M2 simplifies to M1A when the data are simulated from a ZIP model without random effects. At the PH level, we notice identical results for Models 1A and 2A (see Supplementary Table 4A). This is because neither model has a linkage component that jointly describes the risk at both levels. However, M1A yields more unbiased and precise estimates for the county-level model parameters as compared to M2A (see Supplementary Table 5A).

Figures 6A and 7A (see Supplementary Appendix) display the average relative risk estimates obtained from Models 1-4 and Models 1A-2A, respectively. Similar to scenarios 1.1 and 1.2, M3 and M4 recover the clusters of zero county-level risks in Figure 2A, while M1 and M2 recover the elevated county-level risks well. When we compare M1 and M2, M1 recover the simulated county-level risk pattern better than M2. On the other hand, M2 overestimates the clusters of zero risks. Comparing M1A and M2A, M1A returns the simulated county-level risks much better than M2A. Note that the estimated county-level risk pattern from M2 and M1A are almost identical. These results are consistent with the findings in Supplementary Tables 2A and 4A. At PH level, there is a slight difference among Models 1-4 in terms of recovering the simulated risks, while M1A and 2A produces almost identical PH-level risk patterns.

4.1.3 Scenario 2: Simulation from a ZIP Model—The results are displayed in Supplementary Tables 6A–7A. We can see that all the criteria prefer M3 when the data are simulated with 0% and 10% structural zeros. When the structural zeros increase to 25% and 50%, DIC_G selects M1, while DIC_3 and MPL opt for M4. From Supplementary Table 7A, when the percentage of the structural zeros increases, the bias and MSE of the parameters, especially for the intercept and variance-components, obtained from M3 and M4 increase most of the time. When the percentage of structural zeros are 0, 10, and 25, M1 results in more unbiased and precise estimates of the PH-level variance-components (σ_{u2} and σ_{v2}), whereas M2 yields a less biased and more precise estimate for the slope parameter (β_{11}). On the other hand, when the percentage of structural zeros is 50, M2 provides less biased and more precise estimates for the intercept and PH-level variance-components, while M1 produces a less biased and more precise estimate for the county-level variance components.

In terms of recovering the county-level simulated risks, M1 recovers the simulated risks much better than the other models, while M3 and M4 do not recover the simulated risk well, especially the elevated risks when the percentages of structural zeros are 25 and 50. However, when the structural zeros are 0% and 10%, M3 and M4 recover the simulated risks well. Yet, M1 better recovers the elevated risks, For example, in the south west of Georgia. At the PH level, there is a slight difference among the models (see Supplementary Figures 8A, 9A, 10A and 11A).

4.1.4 Scenario 3: Scaling effect—As in scenarios 1 and 2, DIC_3 and MPL pick M3, whereas DIC_G opts for M1 as the county-level best model. In addition, all the criteria select M1 and M3 as the PH-level best model. In contrast to scenarios 1 and 2, all the criteria equally prefer M1 and M3 at the PH level. Comparing between M2 and M4, DIC_3 and MPL

pick M4, while DIC_G chooses M2 at the county level. As in scenario 1, both models perform similarly at the PH level. When we compare M1 to M2, as in scenario 1, M1 outperforms M2 (Table 5).

Table 6 displays the bias and MSE estimates for the model parameters. Unlike in scenario 1, we include the bias and MSE estimates of a_{02} and β_{12} because we are able to fix the values of those PH-level parameters in the simulated PH-level RR. In contrast to scenario 1, M1 does not provide the most unbiased and precise estimates for most of the parameters. Yet, it is doing well for the county-level variance components. On the other hand, M2 offers the best results for the PH-level variance components, while M3 yields the most unbiased estimates of the covariate effects (β_{11} and β_{12}).

Figure 8 depicts the estimated risks, indicating that M1 and M2 recover the county-level simulated risks in Figure 5 well when compared to M3 and M4. Interestingly, at the PH level, the risk patterns from M1 and M3 are similar with the simulated risk patterns in Figure 5. Similar to the simulated risks, the range of the PH-level risks in M1 and M3 is from 0.76 to 3.75, while it is from 0.96 to 3.11 for M2 and M4. At the county-level, the range of the risks for M1 and M2 is from 0.13 to 1.56, whereas it is from 0.05 to 1.34 for M3 and M4. These results show that as in scenarios 1 and 2, M1 and M2 better detect the elevated county-level risks, while M3 and M4 better detect the lower (zero) county-level risks. In contrast to scenarios 1 and 2, the estimated county-level risks are biased because the scale for the simulated county-level risks ranges from 0 to 10, while the estimated one ranges from 0 to 1.6. This is not surprising because the county-level data are simulated from a multinomial distribution, whereas the fitted models assume a county-level Poisson distribution.

4.2 Application to Georgia Skin Cancer Study

Tables 7–8 and Figure 9 show the results obtained from Models 1–4 applied to the Georgia skin cancer study. As in Section 4.1, DIC_3 and MPL suggest that M3 and M4 are the best models at the county level, whereas DIC_G indicates that M1 is the best model at the county level. Moreover, at PH level, all the criteria select M3 as the best model but the difference between M1 and M3 is not significant. When we compare M1 versus M2, we find that M1 outperforms M2, indicating that the shared components in M1 are useful to link the two scale levels simultaneously. From Table 8, we obtain that all the models provide non-significant positive spatial relationship between the daily sunlight and the incidence of skin cancer. However, as in Section 4.1, M3 and M4 yield a negative estimate of the county-level intercept whereas M1 and M2 have positive estimates. We also find that, most of the time, M3 and M4 yield higher estimate of the variance components, especially at the county level, than M1 and M2.

With respect to the RRs, M1 results in estimates that are more consistent across the scales. For example, in northcentral and southeast Georgia, the RR pattern from M1 at both county and PH levels are similar. When we compare the estimated RRs in Figure 9 with the observed SIR in Figure 1, as in Section 4.1, M3 and M4 identify the clusters of zero risks better than M1 and M2. On the other hand, M1 and M2 detect the higher risks better than

M3 and M4. These results show that the zero inflated multiscale models (M1 and M2) tend to overestimate lower risks, while the convolution models (M3 and M4) attempt to underestimate elevated risks.

5 Discussion and Conclusion

Our results indicate that the zero inflated multiscale model with shared components (M1) fits the simulated and the real data better than the zero inflated multiscale model without the shared components (M2) at both the county and PH levels. These findings accentuate the importance of jointly addressing scaling effects and zero inflated problems. We also find that M1 recovers the simulated risk patterns much better than M2. In addition, M1 often offers more unbiased and precise parameter estimates than M2. We also compare the zero inflated multiscale models (M1 and M2) with multiscale models (M3 and M4). We discover that DIC_3 and MPL always prefer the multiscale models over the zero inflated multiscale models, whereas DIC_G often opts for the zero inflated multiscale models. This could be due to DIC_G penalizes more for model complexity than DIC_3 . Another important reason could be that the correlated and uncorrelated random effects could capture the overdispersion due to excessive zeros [33].

To investigate why DIC_3 and MPL opt for the multiscale model instead of the zero inflated multiscale model, we simulate data from a simple Poisson model without random effects but with excessive zeros (scenario 1.3). We fit M1A (zero inflated multiscale model without random effects) and M2A (multiscale model without random effects) to the simulated data sets (see Supplementary Appendix Table 1A). We find that DIC_3 , MPL and DIC_G select M1A as the best model over M2A. However, when we fit the models with the random effects (Models 1-4), DIC_3 and MPL still prefer M3 and M4. This could reflect that the random effects could accommodate the overdispersion due to many zeros.

The zero inflated multiscale models (M1 and M2) often provide more unbiased and precise estimates when compared to the multiscale models (M3 and M4). In scenario 1 where the data are simulated with cluster of zero risks, M3 and M4 yield biased and imprecise estimates for the intercept and some of the county-level variance-components (σ_{u1}). In scenario 2 where we simulate with different percentage of structural zeros, the bias and MSE of the parameter estimates, especially for the intercept, increase as the percentage of structural zero increases (see Supplementary Table 7A). Furthermore, in scenario 3 where we simulate the county-level data from a multinomial distribution conditioning on the simulated PH-level Poisson data, the bias, and MSE for the intercept from M3 and M4 are much larger than that of M1 and M2.

In terms of recovering the simulated risks at the county level, M3 and M4 recover the clusters of zero risks, while M1 and M2 recover the elevated risks well (scenario 1). When the percentages of structural zeros are 25 and 50, M3 and M4 poorly recover the simulated risk patterns, while M1 recover the simulated risk patterns well (scenario 2). When the data are simulated using scenario 3, M1 and M2 recover the simulated risk patterns better than M3 and M4. In all scenarios, there is no significant difference among the models in terms of

recovering the simulated PH-level risks. Although there is a slight difference among the models, all the models recover the PH-level risk patterns well.

Furthermore, we implement the models to investigate whether daily sunlight is spatially related to the incidence of skin cancer in the state of Georgia and find that all the models provide non-significant positive relationship between these two variables at both county and PH levels. M1 provides more consistent risk estimates across scales. As in the simulation study, here also, M3 and M4 better detect the lower (zero) risks, while M1 and M2 identify the elevated risks better than M3 and M4. Moreover, DIC_3 and MPL select M3 and M4 as the best county-level model, whereas DIC_G picks M1 as the best county-level model.

We believe that this paper contributes to the field concerning when and how to implement zero inflated multiscale models to describe risk variation across scales for aggregated data with excessive zeros. The results show zero inflated multiscale models with shared components (M1) are useful to describe risk variations accurately when the percentages of structural zeros are high. Furthermore, they are important to investigate covariate effects across scales. However, several issues require further investigation. First, DIC_3 , MPL and DIC_G measures provide different results. While DIC_G offers expected results for the simulated data sets, DIC_3 and MPL provide unexpected results which appears due to the fact that random effects absorb the overdispersion. Second, we only include a single covariate in the models. This can be extended to include multiple covariate effects. Third, all the models considered here use the data twice because the data at the PH level are an aggregation of the county-level data. To overcome this issue, multilevel models with spatial interaction effects [34–37] could be considered.

Zero inflated multiscale models with shared components address both scaling and zero inflated problems simultaneously. To obtain accurate and consistent risk estimates across scales, we recommend using these models when there are high percentages of structural zeros in the data. Furthermore, these models can be used to detect elevated risks when excessive zeros are present in the data. For identifying areas with zero risks, multiscale models with shared components could be a better choice than the zero inflated multiscale models. However, the zero inflated multiscale models provide more unbiased and precise estimates of the intercept and slope parameters when clusters of zero risks exist in the data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to acknowledge support from the National Institutes of Health via grant R01CA172805. The third and last authors also acknowledge support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy).

References

1. Wong, D. The Modifiable Areal Unit Problem (MAUP). In: Fotheringham, S., Rogerson, P., editors. *The Sage Handbook of Spatial Analysis*. New York: Sage; 2009. p. 108-123.

2. Lawson, AB. Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. 2nd. New York: Chapman and Hall/CRC Press; 2013.
3. Aregay M, Lawson AB, Faes C, Kirby R. Bayesian Multiscale Modeling for Aggregated Disease Mapping Data. *Statistical Methods in Medical Research*. 2015
4. Aregay M, Lawson AB, Faes C, et al. Spatial Mixture Multiscale Modeling for Aggregated Health Data. *Biometrical journal*. 2016; 58:1091–112. [PubMed: 26923178]
5. Aregay M, Lawson AB, Faes C, et al. Multiscale measurement error models for aggregated small area health data. *Statistical Methods in Medical Research*. 2016; 25:1201–23. [PubMed: 27566773]
6. Aregay M, Lawson AB, Faes C, et al. Impact of Income on Small Area Low Birth Weight Incidence Using Multiscale Models. *AIMS Public Health*. 2015; 2(4):667–680. [PubMed: 27398390]
7. Agarwal DK, Gelfand AE, Citron-Pousty S. Zero-inflated models with application to spatial count data. *Environmental and Ecological Statistics*. 2002; 9:341–355.
8. Rathbun SL, Fei S. A spatial zero-inflated Poisson regression model for Oak regeneration. *Environmental and Ecological Statistics*. 2006; 13:409–426.
9. Neelon B, Ghosh P, Loebs PF. A spatial Poisson hurdle model for exploring geographic variation in emergency department visits. *Journal of the Royal Statistical Society: Series A*. 2013; 176:389–413.
10. Neelon, B. Spatial data Analysis for Health Services Research, ch 29. In: Lawson, AB., et al., editors. *Handbook of Spatial Epidemiology*. New York: CRC Press; 2016.
11. Gurmu S, Trivedi PK. Excess Zeros in Count Models for Recreational Trips. *Journal of Business & Economic Statistics*. 1996; 14:469–477.
12. Dalrymple ML, Hudson IL, Ford RPK. Finite Mixture, Zero-inflated Poisson and Hurdle models with application to SIDS. *Computational Statistics & Data Analysis*. 2003; 41:491–504.
13. He H, Tang W, Wang W, Crits-Christoph P. Structural zeroes and zero-inflated models. *Shanghai Archives of Psychiatry*. 2014; 26(4):236–242. [PubMed: 25317011]
14. Boone EL, Stewart-Koster B, Kennard MJ. A hierarchical zero-inflated Poisson regression model for stream fish distribution and abundance. *Environmetrics*. 2012; 23:207–218.
15. Louie MM, Kolaczyk ED. A multiscale method for disease mapping in spatial epidemiology. *Stat Med*. 2006; 25:1287–1306. [PubMed: 16217848]
16. Louie MM, Kolaczyk ED. Multiscale detection of localized anomalous structure in aggregate disease incidence data. *Stat Med*. 2006; 25:787–810. [PubMed: 16453373]
17. Louie MM, Kolaczyk ED. On the covariance properties of certain multiscale spatial processes. *Statist Probab Lett*. 2004; 66:407–416.
18. Kolaczyk ED, Huang H. Multiscale statistical models for hierarchical spatial aggregation. *Geogr Anal*. 2001; 33:95–118.
19. Fonseca TCO, Ferreira MAR. Dynamic Multiscale Spatiotemporal Models for Poisson Data. *JASA*. 2017; 112:215–234.
20. Openshaw S, Taylor PJ. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. *Statistical applications in the spatial sciences*. 1979; 21:127–144.
21. Jelinski DE, Wu J. The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*. 1996; 11:129–140.
22. Gotway CA, Young LJ. Combining incompatible spatial data. *Journal of the American Statistical Association*. 2002; 97:632–648.
23. Gelfand AE, Zhu L, Carlin BP. On the change of support problem for spatio-temporal data. *Biostatistics*. 2001; 2:31–45. [PubMed: 12933555]
24. North America Land Data Assimilation System (NLDAS). Daily Sunlight (insolation) for years 1979–2011 on CDC WONDER Online Database. Centers for Disease Control and Prevention; 2013.
25. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math*. 1991; 43(1):1–20. DOI: 10.1007/bf00116466
26. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*. 1992; 34:1–14.
27. Gelman A. Prior distribution for variance parameters in hierarchical models. *Bayesian Analysis*. 2006; 3:515–533.

28. Celeux G, Forbesy F, Robertz CP, Titterington DM. Deviance Information Criteria for Missing Data Models. *Bayesian Analysis*. 2006; 1:651–674.
29. Spiegelhalter DJ, Best NG, Carlin BP, et al. Bayesian measures of model complexity and fit (with discussion). *J R Stat Soc Series B*. 2002; 64:583–616.
30. Gelman, A., Carlin, JB., Stern, HS., Rubin, DB. *Bayesian Data Analysis*. 2nd. CRC Press; 2004.
31. Bos, C. A Comparison of Marginal Likelihood Computational Methods. Tinbergen Institute; 2002. Discussion Paper TI 2002-084/4
32. Yee, T. Package ‘VGAM’. 2016. p. 811 <http://www.stat.auckland.ac.nz/~yee/VGAM>
33. Neyens T, Lawson AB, Kirby R, et al. Disease mapping of zero-excessive mesothelioma data in Flanders. *Annals of Epidemiology*. 2016; 16:30418–5.
34. Chaix BJ, Merlo J, Chauvin P. Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilization in France. *J Epidemiol Community Health*. 2005; 59:517–526. [PubMed: 15911650]
35. Langford IH, Leyland AH, Rasbash J, Goldstein. Multilevel modelling of the geographical distributions of diseases. *J Roy Stat Soc C-App*. 1999; 48:253–268.
36. Dong, GP., Harris, R., Jones, K., Yu, J. Multilevel Modelling with Spatial Interaction Effects with Application to an Emerging Land Market in Beijing, China. 2015. <http://dx.doi.org/10.1371/journal.pone.0130761>
37. Browne WJ, Goldstein H, Rasbash J. Multiple membership multiple classification (MMMC) models. *Stat Model*. 2001; 1:103–124.

Figure 1. The observed standardized incidence ratio (SIR; top panel) and Average Daily Sunlight (ADS; bottom panel) in thousands of KJ/m^2 at both the county and PH levels.

Figure 2.

The number of persons discharged from non-federal acute-care inpatient facilities for skin cancer in the state of Georgia at both the county (left panel) and Public health (PH) district levels (right panel).

Figure 3.

Scenario 1.1: Simulated relative risks at both the county and PH levels.

Figure 4.
Scenario 1.2: Simulated relative risks at both the county and PH levels

Figure 5.

Scenario 3: Simulated relative risks at both the county and PH levels.

Figure 6.

Scenario 1.1: Averaged relative risks (RR) over 200 simulated data sets at both county (top panel) and public health (PH; bottom panel) levels.

Figure 7.
Scenario 1.2: Averaged relative risks (RR) over 200 simulated data sets at both county (top panel) and public health (PH; bottom panel) levels.

Figure 8.

Scenario 3: Averaged relative risks (RR) over 200 simulated data sets at both county (top panel) and public health (PH; bottom panel) levels.

Figure 9. Georgia skin cancer study: Relative risks (RR) obtained from the models fitted to the data at both county (top panel) and public health (PH; bottom panel) levels.

Table 1
Scenario 1.1: Average DIC_3 , DIC_G and MPL values over 200 simulated data sets.

Models	DIC_3		DIC_G		MPL	
	County	PH district	County	PH district	County	PH district
Model 1	320.95	76.00	362.47	86.64	-148.14	-35.75
Model 2	328.22	78.18	374.57	90.99	-151.36	-36.41
Model 3	298.34	70.83	381.98	78.11	-137.62	-33.82
Model 4	295.38	78.18	378.14	91.08	-134.79	-36.41

Table 2
Scenario 1.1: Average bias and mean square error (MSE) estimates over 200 data sets.

Parameters	M1		M2		M3		M4	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
α_{01}	-0.225	0.082	0.203	0.053	-0.866	0.784	-0.902	0.835
β_{11}	0.041	0.007	0.066	0.011	0.097	0.016	0.104	0.019
σ_{u1}	-0.009	0.014	0.144	0.056	0.308	0.184	1.388	2.085
σ_{v1}	-0.045	0.011	-0.056	0.013	0.231	0.092	0.159	0.073
σ_{u2}	0.601	0.426	0.690	0.534	0.775	0.695	0.696	0.545
σ_{v2}	0.501	0.316	0.596	0.413	0.671	0.544	0.594	0.414

Table 3
Scenario 1.2: Average DIC_3 , DIC_G and MPL values over 200 simulated data sets.

Models	DIC ₃		DIC _G		MPL	
	County	PH district	County	PH district	County	PH district
Model 1	289.17	71.54	322.29	81.79	-133.69	-33.70
Model 2	295.75	75.93	331.80	89.55	-136.99	-35.16
Model 3	272.75	67.25	345.79	74.93	-126.07	-32.05
Model 4	272.89	75.91	354.78	89.57	-124.55	-35.16

Table 4
Scenario 1.2: Average bias and mean square error (MSE) estimates over 200 data sets.

Parameters	M1		M2		M3		M4	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
α_{01}	-0.378	0.188	0.004	0.011	-1.085	1.298	-1.045	1.115
β_{11}	0.030	0.008	0.055	0.011	0.093	0.017	0.084	0.017
σ_{u1}	-0.012	0.016	0.026	0.025	0.353	0.225	1.112	1.386
σ_{v1}	-0.076	0.014	-0.059	0.011	0.169	0.068	0.229	0.106

Table 5

Scenario 3: Average DIC_3 , DIC_G and MPL values over 200 simulated data sets.

Models	DIC ₃		DIC _G		MPL	
	County	PH district	County	PH district	County	PH district
Model 1	172.45	57.98	195.77	63.82	-79.09	-27.49
Model 2	178.56	62.01	207.08	69.23	-81.73	-29.06
Model 3	167.37	57.61	209.91	63.47	-76.52	-27.31
Model 4	171.59	62.02	227.69	69.29	-77.89	-29.06

Table 6
Scenario 3: Average bias and mean square error (MSE) estimates over 200 data sets.

Parameters	M1		M2		M3		M4	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
α_{01}	-1.791	3.302	-1.712	3.011	-2.675	7.274	-2.716	7.520
α_{02}	-0.073	0.088	0.010	0.051	-0.115	0.109	0.009	0.052
β_{11}	-0.060	0.030	-0.026	0.031	-0.007	0.037	0.022	0.039
β_{12}	0.105	0.121	0.434	0.483	0.047	0.110	0.431	0.478
σ_{u1}	0.269	0.137	0.501	0.414	0.537	0.429	0.838	1.007
σ_{v1}	-0.042	0.022	0.133	0.055	0.261	0.126	0.389	0.275
σ_{u2}	0.636	0.562	0.440	0.269	0.713	0.705	0.441	0.276
σ_{v2}	0.276	0.234	0.117	0.089	0.346	0.317	0.123	0.097

Table 7

Georgia skin cancer study: Goodness of fit for Models 1–4.

Models	DIC ₃		DIC _G		MPL	
	County	PH district	County	PH district	County	PH district
Model 1	236.44	67.30	243.61	71.31	−116.0	−32.08
Model 2	239.21	70.79	251.44	74.76	−116.7	−33.72
Model 3	207.60	65.09	279.89	69.20	−94.63	−31.22
Model 4	207.43	68.60	286.20	72.77	−94.44	−32.64

Table 8

Georgia skin cancer study: Posterior summary statistics for the model parameters

Parameters	M1		M2		M3		M4	
	Mean	sd	Mean	sd	Mean	Sd	Mean	sd
α_{01}	0.012	0.350	0.103	0.330	-0.413	0.294	-0.335	0.285
α_{02}	-0.285	0.306	-0.392	0.289	-0.307	0.295	-0.413	0.317
β_{11}	0.128	0.173	0.141	0.163	0.128	0.101	0.135	0.106
β_{12}	0.161	0.193	0.306	0.170	0.169	0.181	0.261	0.190
σ_{u1}	0.482	0.359	0.611	0.396	0.659	0.459	0.763	0.518
σ_{v1}	0.414	0.308	0.359	0.237	0.622	0.330	0.651	0.346
σ_{u2}	0.559	0.393	0.361	0.292	0.579	0.397	0.504	0.368
σ_{v2}	0.253	0.204	0.223	0.212	0.279	0.207	0.235	0.181